



HAL
open science

Data Reduction in multi-hop collection of agriculture data

Joseph McDonnell, Christian Salim, Nathalie Mitton

► **To cite this version:**

Joseph McDonnell, Christian Salim, Nathalie Mitton. Data Reduction in multi-hop collection of agriculture data. International Workshop on Selected Topics in Wireless and Mobile computing (STWiMob), Oct 2024, Paris, France. hal-04684908

HAL Id: hal-04684908

<https://hal.science/hal-04684908>

Submitted on 3 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Data Reduction in multi-hop collection of agriculture data

Joseph McDonnell

Inria

59650 Villeneuve-d'Ascq, France
joseph.mcdonnell@inria.fr

Christian Salim

Junia

Computer Science and Mathematics
59000, Lille, France
christian.salim@junia.com

Nathalie Mitton

Inria

59650 Villeneuve-d'Ascq, France
nathalie.mitton@inria.fr

Abstract—Recent droughts have once again proven the threat of climate change on our food production systems. Building efficient and frugal soil monitoring networks may be part of the solution. Providing real-time data on temperature, humidity and soil moisture can help decision makers to optimize irrigation, fertilization and pest control schedules. Additionally, such networks can be used to raise alert in case of a disease outbreak allowing for a swift response. One of the great challenges to make these tools widely available lays in building frugal and efficient data collection. In part, these networks should have low energy requirements to allow for deployment in remote areas and reduce maintenance costs. Networks with large amounts of transmissions often involve redundancies in the transmitted data, which results in high energy consumption for low utility. For example, the measurement of temperature at a point A and another measurement of temperature at a second point close to A. Using a well-chosen prediction algorithm, estimations of a node's measurements could be an acceptable alternative to transmitting the real values. This work extends our previous work designed for leveraging correlation between two nodes to explore this correlation between multiple nodes over a multi-hop network. Results show that we can reach up to 250 times less data while not losing the information which validates our approach as we can see in the evaluation section.

Index Terms—data reduction, smart agriculture, multi-hop collection.

I. INTRODUCTION

Every field, even on the same farm, has variations in soil, topography, and weather that can affect crop growth. These variations can be caused by factors such as drainage patterns, soil type, and exposure to sunlight. In the past, farmers had to rely on their own observations and experience to make decisions about how to manage their fields. However, this could lead to an excessive use of resources in some areas and under-application in others, resulting in lower crop yields, increased waste and avoidable soil pollution.

Combining human knowledge and experience with information technologies, we can adapt a farming policy to these local differences. This approach fits in a new approach to agriculture called *digital agriculture*, that aims at improving food production whilst preserving the well-being of the environment, the animals, the workers and the consumers involved in the process [1]. To achieve this, farmers can use wireless sensor networks to gather data about these local differences in their land. Sensors measure a wide range of variables, including

temperature, humidity, soil moisture, and crop growth. The resulting data can be analyzed to identify patterns and trends, helping farmers to make more informed decisions about planting, fertilizing, and harvesting. For example, if a farmer knows that one part of the field is drier than the rest, they can adjust the irrigation schedule to provide more water to that area. Similarly, if a farmer knows that one part of the field is more fertile than the rest, they can adjust the fertilization schedule to provide less nutrients to that area.

In addition to improving crop yields, digital agriculture reduces waste of precious resources, such as water and fertilizer, by providing precise information about where and when these resources are needed. Allowing farmers to minimize their environmental impact, making their operations more sustainable and profitable.

Furthermore, wireless sensor networks also help predict potential problems before they become serious issues, like pests or disease outbreaks and take action accordingly.

Overall, digital agriculture and wireless sensor networks are leading the way towards more efficient, sustainable and profitable farming, by providing farmers with the tools they need to make data-driven decisions that can improve crop yields, reduce waste and minimize their environmental impact, while at the same time, providing real-time insights on the field's condition.

In this paper, we mainly focus on reducing the number and size of transmissions within the sensor network. The goal is to reduce to a minimum the number of transmissions whilst retaining a real-time accurate mapping of the land. Each communication step, defined as an emission or reception of data, requires a large amount energy for the sensor and large use of the medium bandwidth. Hence reducing the number of communications increases the available up time of a sensor before a battery needs to be swapped or recharged while improving the capacity of the network. This work represents a real experimentation of our previous works in [2] while testing the algorithms using multi-hop scenarios. Results show an efficiency of more than 50%. The remainder of the paper is divided as follows: Section II presents an overview of the current state of knowledge about wireless sensor networks. Section III formulates the strategy and the used method followed by the experimental results and the conclusion in

Sections IV and V respectively.

II. LITERATURE REVIEW

Data reduction in Wireless Sensor Networks (WSN) has become a focal point of research in recent years, with various techniques employed to optimize energy consumption and minimize data transmission. The primary methods utilized can be categorized into clustering, scheduling, compressive sensing, multi-path multi-channel routing, data correlation, and machine learning, Let's delve into some of these approaches:

Clustering Methods: These techniques target energy conservation in WSNs by identifying similarities in spatio-temporal data [3], [4] and [5]. For instance, in [5], a fuzzy inference system is employed to re-cluster nodes, considering parameters like the average sensed data rate, distance from the sink, and the power of cluster head nodes. Machine learning is then used for data classification based on similarity within each cluster member node, thereby reducing transmission.

Data Correlation and Similarity: Numerous studies have explored correlations and similarities among various data features [3], [6], [7], [8], [9], [10] and [11]. Authors in [11] propose a method for data reduction that relies on assessing the correlation among various variables measured by a single node, using Bayesian inference. They demonstrate its effectiveness by transmitting only atmospheric pressure measurements, which are then used to estimate temperature. This approach reduces the volume of data transmitted by nodes. However, it does not account for correlations between multiple nodes as we consider in our approach.

Authors in [12] went in to details on the levers we can use to reduce the energy consumption of an end device such as the-duty cycle method. However, this raises new problems of synchronizations between devices to make sure receivers will be turned on whenever a node must listen for a new message.

Several techniques exist in the literature that use data prediction to reduce the number of transmissions like in [13] and [14]. [14] In the authors further explore this approach. They introduce intra-node correlation. The reasoning, is that sensors deployed in WSNs will measure multiple variables at a time such as temperature and humidity. These measurements are highly correlated and follow similar trajectories except in the case of an abrupt change in the environment, such as a rainfall. Experimentally their study has proven to work well allowing us to only send the information of sensor A to the base station. The base station can then compute a prediction for second station's measurement based on the correlation coefficient and the mean ratio.

The authors in [15] focused on designing multi-technology networks for precision agriculture. One of the focus points of this thesis was the funneling problem that is encountered by networks when only a select few nodes can communicate with the base station. In this case, an extra load is placed on the nodes who are closest to the sink. This funneling is even worse when only one node is capable of communicating with the sink. Building on the work in [14], they proposes an original

method based on correlations between the transmitted signals. If two end-stations are highly correlated, the base station only needs the data of one node to infer the measurements of the second.

In the next section of this paper, we will be presenting formally the different methods that were implemented and compared in our study.

III. OUR CONTRIBUTION

A. Intra-Node correlation

Each node deployed in a field will be equipped with a number of different sensors to better understand the neighboring environment. At each transmission time step, the node will transmit a packet of information, of size proportional to the number of sensors onboard. Any significant reduction in the size of each individual transmission will reduce the total energy consumption of the end-device. In this section, I will present how the Pearson prediction algorithm can be used to reduce the number of transmissions.

Considering the case of two sensors on board the node, we can define the signals as two series of random variables each following an unspecified distribution. Let's define $X_1, X_2, \dots, X_n \sim \mathcal{X}$ and $Y_1, Y_2, \dots, Y_n \sim \mathcal{Y}$ such that we have n paired observations for both sensors. Following [14], we can define ρ_{XY} and R_{XY} . If ρ_{XY} is above a predefined threshold, we can initialize a prediction process between the node and the sink.

Supposing at step $t + 1$ we decide to predict Y using X , then the node will compute:

$$\hat{y}_{t+1} = y_t + (x_{t+1} - x_t) \times R_{XY} \times \rho_{XY}^2$$

We can then compute the absolute error of the prediction as follows:

$$e_{t+1} = |\hat{y}_{t+1} - y_{t+1}|$$

If e_{t+1} is greater than a predefined threshold, we send both measurements to the sink, otherwise we only send x_{t+1} . On reception, the sink computes \hat{y}_{t+1} using R_{XY} and ρ_{XY} saved in memory. Although, \hat{y}_{t+1} will not be equal to the ground truth we are confident that the residual is smaller than an acceptable threshold.

We can extend this to the case where we have more than two sensors by wisely choosing a base sensor to continue sending to the sink and basing all predictions on this sensor. Formally, if we have N sensors on board the node and each sensor's measurement is of approximately the same size. If we name T the base packet size, and N_{pred} the number of sensors that are accurately predicted. Each packet is sent to the sink using the intra-node data reduction process has a total size in the order of $T \times \frac{N - N_{pred}}{N}$.

In the next section, we will study how inter-node correlation can significantly extend the life-time of nodes in a wireless sensor network.

B. Inter-Node correlation

For the following discussion, consider two neighboring nodes A and B equipped with the same sensors and both relaying their measurements directly to the sink. We aim at using the Pearson prediction process to infer B's measurements using only data relayed by the sensors on A.

Supposing there are m sensors on both nodes we can compute, for each sensor the mean ratio and the correlation coefficient.

$$R_{A_i,B_i}, \rho_{A_i,B_i} \quad i = 1, \dots, m$$

If for sensor i , ρ_{A_i,B_i} is above a certain threshold that entails sufficient correlation we can start the dual prediction process. This means both the sink and sensor B have in memory the parameters required to predict B's measurements.

$$\hat{B}_{i,t+1} = B_{i,t} + (A_{i,t+1} - A_{i,t}) \times R_{A_i,B_i} \times \rho_{A_i,B_i}^2 \quad i = 1, \dots, m \quad (1)$$

Where $A_{i,t}$ refers to the measurement of sensor i on board node A at time-step t .

In this method we suppose that when sensor A relays its measurements to the sink sensor B is in within the receiving range of A's transmissions. Hence, B can use for each sensor i , $A_{i,t+1}$ to compute $\hat{B}_{i,t+1}$.

$$e_{i,t+1} = |\hat{B}_{i,t+1} - B_{i,t+1}| \quad i = 1, \dots, n$$

B can then compare the residuals $e_{i,t+1}$ to the threshold assigned to the error of the prediction of each sensor. If any residual $e_{i,t+1}$ is greater than the acceptable threshold it will transmit $B_{i,t+1}$ to A who will in turn relay the ground truth to the sink. If the residual is smaller than the acceptable threshold the sink will compute $\hat{B}_{i,t+1}$ according to equation 1. In both cases, we are guaranteed that the sink records all measurements within a certain degree of accuracy to the ground truth.

At any time step where the sink correctly infers B's measurements, the total number of transmissions across our theoretical network are halved. Activating the emitter for a transmission is the most energy intensive task for a node in a wireless sensor network. Foregoing this step significantly extends the lifetime of the node. There are however additional elements that need to be considered to make this algorithm fruitful. Extending the lifetime of a single node is not very effective. An additional load balancing algorithm should be implemented to guarantee that nodes are switched off in turns to extend the lifetime of the whole network without losing any data in the process. Secondly, we need a set-up where the node used for sending data is in communication range with both the sink and the neighboring node.

In the next section we will study the application of the Pearson prediction process for a multi-hop network, a more commonly used routing set-up.

C. Multi-Hop Network

When all the nodes of the network are not in communication range with the sink, we resort to relaying messages through

a path of nodes until reaching the sink. This strategy is also known as multi-hop routing. Transmission loads are unbalanced between nodes closest to sink with a large amount of transmissions and nodes without. As the message travels through the network the total packet size to send increases as each relay node adds on his own measurements. In this section, we will show how the Pearson Prediction Algorithm can significantly reduce the total load of the network and balance the transmission workload between all nodes.

Consider a network with n nodes labeled N_1, N_2, \dots, N_n where node 1 is nearest to the sink, node 2 communicates with the sink by relaying through N_1 , node 3 relays through N_2 and N_1 , and so on. In this case, if we consider N_0 to be sink, we suppose that any node $i = 1, 2, \dots, n$ is in communication range with nodes N_{i-1} and N_{i+1} . Additionally, we suppose all nodes in the network are identical in terms of sensors and acquisition frequencies.

Without the prediction algorithm, node N_n would send his measurements to N_{n-1} and so on until reaching node N_1 who will send the measurements of the entire network to the sink. The Pearson prediction algorithm in this scenario aims at reducing the bottle-neck for nodes at the end of the network. Fundamentally, this data reduction algorithm relies on high correlation between nodes for the sink to predict measurements using only the real values sent in from N_1 . To achieve this, the sink propagates through the network by predicting the value of a node using the predicted value of its successor.

Initialization of this routing protocol starts by computing all the parameters for each step of the prediction process. Over a set period of time, using past values, for each node $i = 1, \dots, n$ and for each sensor on board j , the sink computes $R_{N_{i-1}, N_i, j}$ and $\rho_{N_{i-1}, N_i, j}$. The sink will then send out these parameters to the network such that each node has them in memory. The idea is such that each node has the necessary parameters to compare the predicted values of the sink with the real measurements the node observed.

Once the training is completed, we can begin with the energy saving strategy. At each time-step, N_1 sends its measurements to the sink and a recursive process begins. N_2 being in communication range with N_1 also receives the measurements. Having $R_{N_{i-1}, N_i, j}$ and $\rho_{N_{i-1}, N_i, j}$ in memory it can compute for each sensor j , the predicted value based on N_1 's measurements.

$$\hat{N}_{2,j,t+1} = N_{2,j,t} + (N_{2,j,t+1} - N_{1,j,t}) \times R_{N_{1,j}, N_{2,j}} \times \rho_{N_{1,j}, N_{2,j}}^2 \quad j = 1, 2,$$

Where $N_{2,j,t}$ designates the measurement at time-step t , of sensor j on board node N_2 .

Similarly to inter-node correlation N_2 can compute the residuals of each sensor and check if they are inferior to a certain threshold. Based on this result of the residual check, if the prediction is acceptable in terms of accuracy it communicates to N_3 the prediction $\hat{N}_{2,j,t+1}$. Otherwise it sends the ground truth measured by its sensors. This step is repeated at each node until reaching the end of the network. Any node with an incorrect prediction will then send its ground

truth to the sink. The correction step starts at the end of the network such that each node with an incorrect prediction can add on its own ground truth as the packet travels along the route.

Considering a time-step where all the predictions are correct, a single-node will only transmit when sending its own predicted values, except for the node N_n at the far end of the network. This means for any node i in the network, the saved size of the packet at each time-step is equal to $n + 1 - i$. In the case where an error is detected we can choose between different error correction strategies. The baseline protocol would be to transmit the ground truth to the succeeding node who would relay it to the sink as suggested above. This means in case of an error at node N_i an additional i transmissions will be performed to correct the error. In the case where an error may entail an outlier that requires quick handling by the network operator. The node could use a large amount of energy to transmit the outlier immediately to the sink.

IV. EVALUATION

A. Dataset Description

In this paper, the data in the studied dataset is collected from 42 nodes dispatched over 42ha in cook agronomy farm state of Washington city of Pullman. Each node is immersed 90 cm underground. On each node we have a sensor on the ground level and sensors at each underground level 30cm, 60cm and 90cm depth. On each depth, the sensors collect the soil temperature and humidity alongside other soil measurements. Data is collected each hour, the network has been in operation since 2009. The purpose of this paper, as mentioned before, is to compute the correlation between data from different depths to check whether we can predict the values from each other instead of collecting them. In this case, the energy consumption related to transmission and collection is reduced.

B. Intra-Node correlation

As explained in III-A, each individual node of a network is equipped with multiple sensors to accurately measure and describe its environment. In the case of agriculture, such sensors will measure the parameters of the surrounding atmosphere such as the temperature, humidity levels or wind speed. It could also be recording the same parameters at different soil depths. In the following section, we will study through experiments the effectiveness of the Pearson algorithm in the case of intra-node prediction. Here, we examine the precision of the Pearson prediction algorithm in the case of a node measuring simultaneously temperature and humidity at 30cm depth inside the soil. We studied the performance of the algorithm for each different season. For the experiment, we used a month of past data to compute the mean ratio and the correlation coefficient. We then iterated over the data of the following year to check the accuracy of our method.

Figure 1, presents the results of our experiment. The left figure presents the accuracy when predicting temperature using the measured humidity percentage. Conversely, the rightmost figure presents the accuracy when predicting the humidity

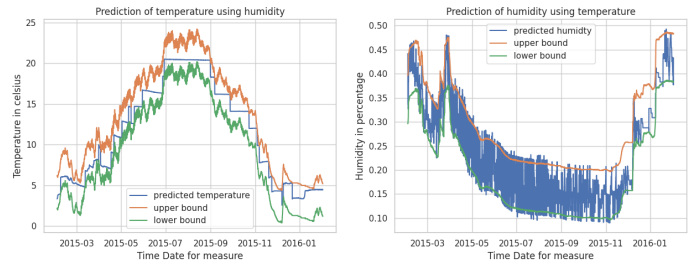


Fig. 1: Results of experiment on node

percentage using the temperature in Celsius and the Pearson prediction algorithm. The upper and lower bound drawn respectively in orange and green represent the tolerance range for the accuracy of the prediction. For all the experiments that follow in this report we use a tolerance range of 2°C for temperature and 5% for humidity.

Basing our temperature prediction off humidity is highly effective resulting in only 12 predicted values outside of the tolerated error range. On the other hand, predicting the humidity percentage based on the temperature resulted in 161 errors. Additionally, when predicting the temperature our final result appears to have a smooth ladder like appearance given the stable evolution of the humidity percentage. In contrast, the cyclical evolution of the temperature between day and night periods returns a prediction with a large amount of variations. Even though predicting humidity based off temperature returns a decrease of 48% in the total size of transmissions the accuracy has a higher volatility. Meanwhile, predicting temperature using humidity returns a decrease of 50% in the total size of transmissions given the predefined tolerance parameters.

The nodes in the available dataset recorded both temperature and humidity at different levels of depth beneath the ground.



Fig. 2: Results of the prediction process at different depths

The goal of this experiment is to demonstrate the effectiveness of predicting the measurements of sensors at 60 and 90cm of depth using the sensor at 30cm as a baseline. The set up follows the same steps as the previous experiment and figure 2 presents the results in a similar manner. We observe that for both sensors and both depths the algorithm is highly effective. Indeed, over 5 months with hourly measurements the worst performing sensor only required 18 corrections.

This experiment suggests that applying the intra-node Pearson prediction algorithm could decrease by 66% the total packet size to transmit to the sink at each time-step.

These two experiments, demonstrate the effectiveness of predicting temperature at 30cm depth using the humidity percentage with the Pearson prediction algorithm. It also shows impressive performance at predicting the same measurements at different depths.

C. Inter-node Correlation

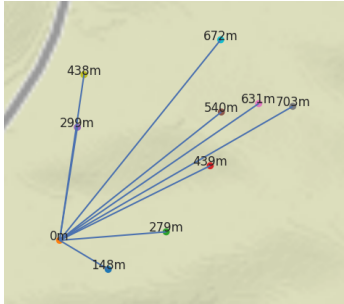


Fig. 3: Map of nodes used for experiment

Digital agriculture increasingly relies on wireless sensor networks, which often results in redundant information being sent to the sink by neighboring sensors. In this study, we investigate the performance of an algorithm for predicting measurements between nodes in such networks. Specifically, we aim to explore how the algorithm handles the issue of redundancy, and provide insights into its effectiveness in optimizing the use of network resources.

Our experimental setup is centered around the Cook experimental farm dataset, which consists of 9 nodes. The sensor layout is depicted in figure 3. Our study focuses on the accuracy of the nodes in predicting measurements for distances ranging between 148m and 703m, using the signal of the node labeled *CAF129* as a reference.

To evaluate the accuracy of our algorithm, we used the first month of measurements for parameter initialization and the subsequent months for performance evaluation. Despite the small distances between sensors, the correlation coefficient between all pairs of nodes was greater than 99%. Our experiments revealed that for 8 different nodes, the algorithm made no more than 6 errors when predicting temperature or humidity over a year, with some sensors never exceeding the error threshold. We observed the mean absolute error of each sensor as a function of distance to the master node. Notably, our algorithm achieved high precision in predicting humidity, with a mean absolute error that never exceeded half of the considered threshold. The same can be said for the prediction of temperature. This experiment demonstrates that for nodes in a range of 700m the suggested algorithm is highly accurate.

D. Multi-Hop Network

This final experiment regroups both of the previous sections and allows us to analyse the performance of the prediction

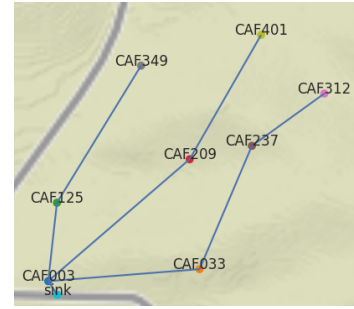


Fig. 4: Graph representation of the experimental multi-hop network

process in a realistic network. For this experiment we will be using the measurements of a subset of eight nodes between March 2013 and April 2014. Figure 4, gives a geographically accurate representation of the nodes we will be using set on Cook farm in Washington State, USA. By design, the network covers the terrain available with variations in distance between nodes and path construction. All the data acquired by the nodes is relayed to the sink following the path in blue. Arises the funneling problem for node *CAF003* because it is the only node chosen to be in communication range with the sink. Here we are primarily interested in studying the accuracy of the prediction process as the number of consecutive predictions increases and the factor by which the number and size of the transmissions in the network are reduced. For the simulation we use a month of measurements for training data and analyze our results on the performance of the algorithm throughout the following year.

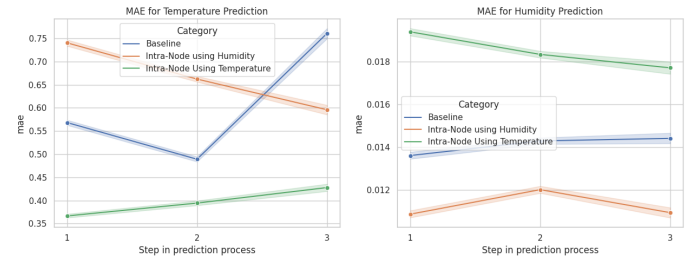


Fig. 5: Mean Average Error of Prediction Process

Figure 5, presents the mean absolute error for the prediction process at different steps in the network. When the step in the prediction process is equal to 1 this means we are only one step behind in the network from the node that sent the ground truth to the sink. Similarly, step 2 and 3 refer to nodes for which predictions are respectively 1 or 2 steps behind in the prediction process from the ground truth. Each marker refers to the mean absolute error and the tails of the plot represent the 95% confidence intervals. The first take back from this experiment is the relatively high-accuracy even for the nodes further away from the sink. It is important to keep in mind that an accuracy that is considered as acceptable for temperature is within 2°C for the temperature and 5% for humidity. Here we are firmly withing these error margins for every tested methods

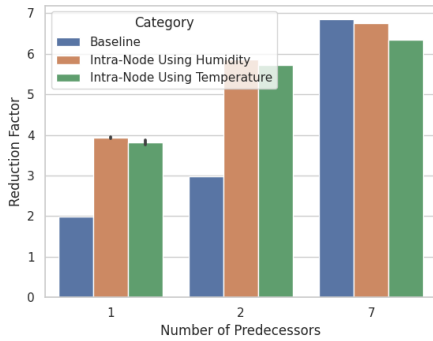


Fig. 6: Reduction factor of proposed method

at each step of the network. We are also able to compare between the baseline model that does not integrate intra-node predictions, the model in orange that predicts humidity using the prediction of temperature and the model in green that predicts the opposite. The graph offers a good representation of the trade-off when implementing this model. The variable on which the intra-node prediction is based-from is slightly more accurate in both cases than the baseline model without intra-node prediction. Conversely, the variable predicted using intra-node prediction is slightly less accurate compared to the baseline.

Reducing the total energy consumption of the network is the end goal of this algorithm. As shown above reduce this energy consumption entails reducing the size and number of transmissions in the network. Figure 6, shows the reduction factor compared to the scenario without an energy reduction method. We base our representation on the number of predecessors of the current node in the network. For example, the node with 7 predecessors is named CAF003 and has the highest load because it must relay all the messages of the network to the sink. This figure proves the promising results of this new method and shows that, in this experiment, the energy consumption related to transmission of a single node is at a minimum divided by two. Intra-Node correlation also proves to be key in reducing the number of transmissions. Adding intra-node correlation multiplies the reduction factor nearly by two. The single drawback of adding intra-node correlation is the higher number of erroneous predictions that need to be corrected for. Most importantly, the node at the end of the network benefits from the most important reduction factor that is in average around 250. Indeed, since nodes without predecessors do not have any messages to relay it only communicated when its own predicted values are outside of the acceptable error range. Given the proven accuracy of this method, these transmissions are few and far apart meaning these nodes benefit the most from this data reduction process.

V. CONCLUSION

In summary, this paper investigates the potential utility of the Pearson prediction process in minimizing energy consumption in a wireless sensor network. After describing the

algorithm and its potential applications, we conducted experiments using the Cook agronomy farm dataset. The results were highly encouraging, with significant reductions observed in transmission size and number when operating in a multi-hop network configuration. These results suggest that this simple algorithm is very effective in this particular scenario. Additionally, a well designed routing algorithm built around the Pearson prediction process may lead to a significant reduction in data consumption for wireless sensor networks.

REFERENCES

- [1] "Agriculture et numérique."
- [2] C. Salim and N. Mitton, "K-predictions based data reduction approach in wsn for smart agriculture," *Computing*, vol. 103, 2021. [Online]. Available: <https://doi.org/10.1007/s00607-020-00864-z>
- [3] G. B. Tayeh, A. Makhoul, C. Perera, and J. Demerjian, "A spatial-temporal correlation approach for data reduction in cluster-based sensor networks," 2019.
- [4] M. Wu, L. Tan, and N. Xiong, "Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications," *Information Sciences*, vol. 329, pp. 800 – 818, 2016, special issue on Discovery Science.
- [5] S. Radhika and P. Rangarajan, "On improving the lifespan of wireless sensor networks with fuzzy based clustering and machine learning based data reduction," *Applied Soft Computing*, vol. 83, 2019.
- [6] A. Ghaddar, T. Razafindralambo, I. Simplot-Ryl, D. Simplot-Ryl, S. Tawbi, and A. Hijazi, "Investigating Data Similarity and Estimation Through Spatio-Temporal Correlation to Enhance Energy Efficiency in WSNs," *Ad Hoc & Sensor Wireless Networks*, vol. 16, no. 4, pp. 273–295, Dec. 2012.
- [7] C. Razafimandimby, V. Loscri, A. Maria Vegni, D. Aourir, and A. Neri, "A Bayesian approach for an efficient data reduction in IoT," in *Int. Conf. on Interoperability in IoT (InterIoT)*, Nov. 2017.
- [8] C. Razafimandimby, V. Loscri, A. M. Vegni, and A. Neri, "Efficient bayesian communication approach for smart agriculture applications," in *IEEE Vehicular Technology Conf. (VTC-Fall)*, Sep. 2017.
- [9] M. Azaza, C. Tanougast, E. Fabrizio, and A. Mami, "Smart greenhouse fuzzy logic based control system enhanced with wireless data monitoring," *ISA Transactions*, vol. 61, pp. 297 – 307, 2016.
- [10] A. Ghaddar, T. Razafindralambo, I. Simplot-Ryl, S. Tawbi, and A. Hijazi, "Algorithm for data similarity measurements to reduce data redundancy in wireless sensor networks," in *A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, June 2010.
- [11] C. Razafimandimby, V. Loscri, A. Maria Vegni, D. Aourir, and A. Neri, "A Bayesian approach for an efficient data reduction in IoT," in *InterIoT 2017 - 3rd EAI International Conference on Interoperability in IoT*, Valencia, Spain, Nov. 2017. [Online]. Available: <https://hal.inria.fr/hal-01620373>
- [12] S. M. Chowdhury and A. Hossain, "Different Energy Saving Schemes in Wireless Sensor Networks: A Survey," *Wireless Personal Communications*, vol. 114, no. 3, pp. 2043–2062, Oct. 2020. [Online]. Available: <https://doi.org/10.1007/s11277-020-07461-5>
- [13] G. B. Tayeh, A. Makhoul, D. Laiymani, and J. Demerjian, "A distributed real-time data prediction and adaptive sensing approach for wireless sensor networks," *Pervasive and Mobile Computing*, vol. 49, pp. 62–75, Sep. 2018. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1574119217305886>
- [14] C. Salim and N. Mitton, "K-predictions based data reduction approach in WSN for smart agriculture," *Computing*, vol. 103, no. 3, pp. 509–532, Mar. 2021. [Online]. Available: <http://link.springer.com/10.1007/s00607-020-00864-z>
- [15] I. Guérin-Lassous, A.-L. Beylot, T. Noël, M. D. de Amorim, N. Mitton, and M. Bruniaux, "Communication sans fil polymorphe pour l'agriculture connectée."